DOCUMENT RESUME

ED 088 750 SO 007 099

AUTHOR Taylor, Bob L.
TITLE Use of the Model at the State and Local Levels.
PUB DATE [73]
NOTE 15p.; Paper presented at the National Council for Social Studies (San Francisco, November, 1973)

EDRS PRICE MF-$0.75 HC-$1.50
DESCRIPTORS Academic Achievement; *Achievement Rating; Comparative Analysis; Course Evaluation; Criterion Referenced Tests; Educational Accountability; *Educational Assessment; *Evaluation Methods; *Models; National Norms; *National Surveys; Norm Referenced Tests; Program Evaluation; School Districts; School Surveys; State Programs; Statewide Planning; Student Evaluation
IDENTIFIERS *National Assessment

ABSTRACT

The National Assessment, a census-like study to collect information concerning the educational attainment of Americans, is being adapted in many states with the results of the adaptations being used for decision making by state agencies or by teachers and administrators. The characteristics of the adaptations follow patterns related to this distinction of intended user. From among those states which have adapted the National Assessment as a model for assessment programs, this paper discusses the state-level programs in Maine, Connecticut, Texas, and Colorado and the district-level programs in Nebraska and Maryland. Although certain characteristics denote a good assessment program, the many possible variations open the model to misuse. Appropriate use of the national model can promote curriculum improvement and yield valuable information for decision making; for example, the necessity of determining behavioral objectives can lead to sharpened perceptions of educational aims, but the exercise can also narrow perspectives. Assessment data too can be misinterpreted and misapplied. An optimum use of the model is for accountability when applied to a total organization, such as a school. (JH)

ED 088750

U.S. DEPARTMENT OF HEALTH, EDUCATION & WELFARE
NATIONAL INSTITUTE OF EDUCATION
THIS DOCUMENT HAS BEEN REPRODUCED EXACTLY AS RECEIVED FROM THE PERSON OR ORGANIZATION ORIGINATING IT. POINTS OF VIEW OR OPINIONS STATED DO NOT NECESSARILY REPRESENT OFFICIAL NATIONAL INSTITUTE OF EDUCATION POSITION OR POLICY

USE OF THE MODEL AT STATE AND LOCAL LEVELS

by

Bob L. Taylor
University of Colorado

SP 007 099

National Assessment is a census-like study to collect information concerning the educational attainments of Americans. In planning for the collection of this census-like data, a model for gathering information about the achievement levels of students and young adults was developed. A number of states have found adaptations of the model useful in conducting state assessments in which desirable learning outcomes are identified and the status of learners with respect to these outcomes is determined.

State assessment is a rapidly developing movement. At this writing, all of the states have assessment activities either in operation, in a developmental process, or in a planning stage.[1] While the statewide assessment programs have many similarities, they break down into two basic types of programs on the question, "Who gets to use the results?" The divisions are those states for which data are collected for decision making by state agencies and those states for which data are collected for decision making by teachers and administrators.

A number of characteristics are appearing in these assessment programs. In about a third of the states, the programs were mandated by the state legislatures, and the results of the assessments

---

[1] State Educational Assessment Programs, 1973 Revision, Joan S. Beers and Paul B. Campbell, "Statewide Educational Assessment," Educational Testing Service, Princeton, New Jersey, 1973, p. 1.

are to be reported back to the state legislatures. In a few of the states, the data are to be used for PPBS (Planning, Programming, and Budgeting Systems). In about half of the states where the assessment data are being used to make state-level decisions, state and Federal funds will be allocated based on the results. Participation in assessment is required by law in about a fifth of the states. In the states where the assessment data are being used to make state-level decisions, samples rather than all students are being assessed, while in the local-level, decision-making states all students in the target populations are being assessed. Criterion-referenced instruments are very common with the states where the data are being used for state-level decisions, but the states collecting information for local decision making are favoring norm-referenced instruments. Finally, no dominant funding pattern has evolved in either of the two groups of states.[2]

State Adaptations of the Model

In the assessment of Citizenship education, Maine made an extensive application of the National Assessment model and carefully duplicated it so that comparable data were collected at the state level.[3] Maine's first cycle of the ten subject matter areas of National Assessment (Art, Career and Occupational Development, Citizenship, Literature, Mathematics, Music, Reading, Science, Social Studies, and Writing) is to be completed by scheduling two of these areas each year for five years. Citizenship and Writing were the first subject areas to be assessed.

---

[2] Ibid., pp. 2-3.

[3] Maine Assessment of Educational Progress: Methodology (Report 5), Department of Educational and Cultural Services, Augusta, Maine, 1972.

Based on the results of a previous study of objectives for education in Maine, two review committees decided to accept the National Assessment objectives as being closely related to the Maine objectives. Maine selected the 17-year-old population of in-school students for its first assessment. A sample of 2,000 17-year-old students was used to represent the approximately 17,000 17-year-old students in the State. The State was divided into four geographical regions. As in National Assessment, school buildings were randomly selected from the geographic regions, and students were then randomly selected from the buildings. Packages were developed with exercises taken from the two subject areas. The available, released exercises from National Assessment were carefully examined to see if they reflected objectives valid for Maine and to see if some could be modified, where needed, to be administered in group sessions using the paced-tape method while still retaining a high degree of comparability to the National Assessment individually administered exercises. The packages were made up of 23 Citizenship and seven Writing exercises, plus a 23-item Student Questionnaire. The exercise format was kept virtually identical to the one used in National Assessment. Trained administrators were sent out to administer the exercises, and the exercises were scored according to National Assessment procedures. On data reporting and analysis, there was the census-like reporting of the performance of the Maine students plus comparisons of the Maine results with appropriate National Assessment data.

In summary, the Maine Assessment duplicated the National Assessment procedure as completely as possible. With minor exceptions, the same objectives were used for Citizenship. The same

sampling design was used with adaptations to a smaller geographical area and population. The exercises were for the most part taken from those released by National Assessment, and they were organized into packages similar to those used by National Assessment. The administration and scoring of the exercises were conducted in the same manner as National Assessment had used. Since the same private contractors were used by Maine as were used by National Assessment, the duplication was complete wherever possible. The reporting and data analysis were similar, and the data did provide the opportunity to compare the results in Maine with the results from National Assessment.

Here, the model was very carefully duplicated at the state level. The big question which comes to mind after studying the Maine Citizenship report is, "Aren't the National Assessment data being treated here as some kind of a national norm against which the performances of 17-year-old students in Maine were being compared?" Of course, this use of National Assessment data had been questioned from the start of the proposal for an assessment at the national level. Now, Maine has provided the opportunity to study the effects of this use of the data on the educational system of a state.

Another state which carefully followed the model was Connecticut.[4] Here, an assessment was first conducted in Reading. To permit comparisons, the Connecticut program used available instruments and applicable procedures developed by National Assessment which were adapted to the requirements of the local situation.

---

[4] Report on the Assessment of Reading Skills of Connecticut Public School Students, Institute for the Study of Inquiring Systems, Philadelphia, Pa., and Department of Education, Hartford, Conn., 1972.

Connecticut's Reading objectives were matched to the Reading objectives of National Assessment. Approximately 220 reading exercises from National Assessment were used in producing the packages used in the Connecticut assessment. Exercises were selected to represent all of Connecticut's Reading objectives. The age groups assessed were 9, 13, and 17. As with the National Assessment packages, tape-recorded instructions were used. The sampling design was a multi-staged design duplicating with few exceptions the National Assessment design. As with National Assessment, a group of administrators for the packages was recruited and trained.

This was another example of careful duplication of the National Assessment model down to using the same objectives and exercises. Again, there was the use of the National Assessment results as norms to which the Connecticut results were compared.

The Texas Needs Assessment used the model for the development of their assessment in Mathematics at the sixth-grade level.[5] However, while using ideas from the model, they broke with it in a number of places. The Texas people were concerned that the assessment would yield information which would be useful to teachers in their classroom instruction of students. From a pilot study, it was decided to use a criterion-referenced reading test and to work with grade levels instead of age groups of students. They worked with the sixth grade, and the tests were administered by the staff of each school which participated in the assessment. The objectives were chosen from the major skill areas treated in the state-adopted textbooks. Regional location and community size were taken into

---

[5] Sixth-Grade Mathematics: A Needs Assessment Report, Texas Education Agency, Austin, Texas, 1972.

consideration in selecting the sample. Approximately 10 per cent of the Texas schools teaching at the sixth-grade level administered tests, and approximately 10 per cent of the pupils being taught at the sixth-grade level were included in the sample. Reports were given to teachers on the performance of their individual students. Also, there was a school report on the performance of the students for each school and a report on each of the classes in the school. Comparisons were made on the basis of sex, race, and size of community.

The Colorado Needs Assessment, while using the model, made an even greater break with it.[6] Its objectives were based on a state study of educational goals, and the educational goals were restated in terms of performance objectives. Following the model, objective-referenced exercises were written. A sampling design was used and the student responses were analyzed. In this assessment, classroom teachers were involved in the writing and refinement of the behavioral objectives. Objective-referenced exercises were written for nine subject areas. The exercises were administered to a sample of 30,000 Colorado students. A stratified random sampling procedure was used to select a sample of school districts of the State. Then schools were selected at random from the districts chosen. Finally, classes in school buildings were randomly chosen for testing. The samples were representative of all Colorado students in grades 3, 6, 9, and 12. A group of proctors was hired and trained to administer the exercises, and the exercises were scored by

---

[6]Helper, John W., An Assessment of Learner Needs in Colorado, Colorado Department of Education, Denver, Colo., 1972.

computer. The data were analyzed on a statewide and district basis, and the results were broken down by subgroups, e.g. boys, girls, urban, rural.

As pointed out earlier, states are rapidly moving into the assessment field. Some are reproducing the National Assessment model at the state level, and others are developing variations of the model. The more crude efforts have resulted in endless pages of raw percentages without any explanation of the results. Based on a survey of state assessment programs, Beers and Campbell identified several of the problems which are common to these state programs.[7] Naturally, a shortage of money and staff were the most frequently mentioned problems, for it is a fact that many states have moved into this area without providing adequate funds for a realistic assessment program. Also, teacher resistance to assessment and negative public attitude toward outside testing were problems mentioned. Test results have been misused in the past, such as the firing of teachers based on incorrect interpretation of test results. Also, test developers have been guilty of violating the privacy of students through questions which transgressed the examinee's human and legal rights. A third problem area has been with the utilization and dissemination of results. Some school officials do not understand the results. In some situations, there has been hostility to the results. Some officials have ignored results in making decisions. Finally, results have frequently not reached the right people in a useable form.

---

[7] State Educational Assessment Programs, Beers and Campbell, op. cit., p. 3.

Use of the Model at the District Level

To date, there have been a limited number of efforts reported on the use of the National Assessment model at the district level. Three such assessments on which some data have been released are being conducted in Lincoln, Nebraska; San Bernardino, California; and Montgomery County, Maryland.

In Lincoln, Nebraska, the exercises released by National Assessment in Citizenship and Writing were used in a local assessment which yielded data comparable to National Assessment data.[8,9] A group of supervisory personnel from the central office identified the Citizenship objectives which were applicable to the Lincoln schools. Then the released National Assessment exercises were selected which were applicable to these Lincoln objectives. Also, the National Assessment model was followed in selecting a random sample of 13 year olds from the Lincoln junior high schools. In addition, a sample of in-school, 17 year olds was tested on some of the writing exercises. The administration of the exercises was carried out by a group of specially trained administrators, and the tape-paced method was used in presenting the exercises to the students. Scoring followed the National Assessment procedures, and in reporting the results comparisons were made to National Assessment data with special attention given to comparable subgroups such as cities of similar size and the same geographic region.

The San Bernardino City Schools developed a criterion-referenced assessment model of student progress which was based on the National

---

[8]"Weekly Focus," Lincoln Public Schools, Lincoln, Nebraska, February 12-19, 1973, p. 3.

[9]Brandt, Ronald, Associate Superintendent for Instruction, Report on Assessment Results to Board of Education, Lincoln Public Schools, Lincoln, Nebraska, Spring, 1973.

Assessment model.[10,11] This model involved local teachers, students, and laymen in a good setting; eight educational goals were identified through the efforts of workshops involving teachers, students, and patrons. A Curriculum Task Force composed of 20 teachers wrote behavioral objectives for the goals to be appropriate for grades 3, 5, 9, and 12. National Assessment consultants assisted the teachers in developing exercises to assess the stated objectives at these grade levels. Also, the National Assessment consultants helped to design a sampling procedure to provide district-wide representation. The exercises were organized into test batteries for each grade level. The Teacher Task Force administered and scored the tests. This is a break from the National Assessment practice of using specially trained exercise administrators. The results were tabulated in terms of percentage of students meeting stated behavioral objectives. While the National Assessment model was followed in many ways, such as use of behavioral objectives, criterion-referenced assessment instruments, and sampling of target populations, the assessment was designed for application at the local level, and it was planned, developed, and carried out by local personnel.

The Montgomery County Schools, Maryland, developed a program for assessing 13- and 17-year-old students.[12] In this assessment,

---

[10]Bonney, Lewis A., "Application of the National Assessment of Educational Progress Philosophy in San Bernardino City Unified School District," Unpublished Paper, San Bernardino City Unified School District, San Bernardino, California.

[11]Special Curriculum Task Force, "Report on Student Performance," Office of Instructional Services and Research and Development Office, San Bernardino City Unified School District, San Bernardino, California, June, 1972.

[12]Bayless, David L., Ralph E. Folsom, and Louise H. Lewis, "Sample Design for Assessing Montgomery County Public Schools 13- and 17-Year-Old Pupils Using the NAEP Model," National Assessment of Educational Progress and Educational Commission of the States, Denver, Colorado, January, 1973.

the released National Assessment exercises for Writing were used. These were administered in two group-package sessions to samples of 13- and 17-year-old students. The results for Montgomery County students were compared to the results from the nationwide samplings of 13 and 17 year olds by National Assessment. One of the variations in the Montgomery County sampling design was stratification by I.Q. and grade level. The purpose was to spread the sample across the grade by school-I.Q. groups; however, these sampling groups were not used as reporting units. Each age group (13 and 17) was stratified by I.Q. groups (low and nonlow) and by grade levels.

Adaptation of the Model

In the above discussion, it is evident that there will be as many adaptations of the model as there are local and state units conducting assessments. Probably, there is no specific assessment model which is the best; hence, there is no model that should be applied without modification in any and all situations. Nevertheless, there are principles of good assessment which should be applied in developing or adapting a model for local assessment purposes. Listed below are some characteristics which should be found in a good assessment program.[13]

1. The program has clearly defined goals that apply to a particular audience or audiences.

2. The program has a realistic number of goals which are attainable under the existing assessing conditions.

3. The program has established priorities among its goals and places its major efforts on its major goals.

4. The program has been designed to gather information considered to be important in education.

---

[13]Womer, Frank B., Developing a Large Scale Assessment Program, Cooperative Accountability Project, Denver, Colo., 1973, p. 89.

5. The program has specific objectives which it is striving to attain.

6. The program has been designed to provide results at a useable level of accuracy.

7. The program has used data-gathering instruments which measure the objectives of the assessment.

8. The program has collected data in such a manner as to introduce a minimum of error in the results.

9. The program has scored and processed data in an accurate manner.

10. The program has used analytic techniques that provide the data breakdowns needed by decision makers.

11. The program has reported results in a manner useable by its audience.

12. The program has provided help in the interpretation of results and assistance in their implementation.

13. The program has provided for the active involvement of groups of persons from all of the major audiences for the assessment results.

Implications of the Model

The assessment model has potential for promoting curriculum development. This is especially true when it is applied to state or local situations in the manner used in Colorado and San Bernardino. In these two situations, objectives were developed which specifically applied to the local situation. The statement of well written objectives in behavioral terms may sharpen the purposes of instruction. Through the experience of writing behavioral objectives, the curriculum worker gains a much clearer perception of his task; hence, this practice may have a beneficial impact on curriculum work. On the other hand, the use of behavioral objectives has not always been a positive influence. The objectives may zero in on easily defined behaviors which lack scope and significance. They may produce tunnel

vision, and put stress on the inconsequential and trivial. In an effort to be specific and to define the exact behaviors desired, the larger perspective may be lost.

Again, the development of exercises from the identified behavioral objectives may have a positive influence on curriculum. The kind of new, innovative exercises which have been developed by National Assessment may have a very positive influence on what is being taught and how it is being taught. Teachers both in reviewing exercises which have been used in National Assessment and in writing exercises for local assessments may be influenced in their selection of both content and methods by their knowledge of these assessment exercises. Material not relevant to the objectives of the course may be dropped, and methodologies promoting the kind of skills needed in the assessment exercises may be introduced.

On the other hand, the results may be less desirable. If in local and state situations the dictates of finances or the lack of leadership results in the use of poorly written, machine-scored, multiple-choice exercises, the results may be very negative. Teachers may feel pressured to stress rote learning of facts in order to prepare their students for poorly written examinations. Hence, poorly written exercises may keep irrelevant material in the curriculum and limit curriculum innovation and development. The quality of the exercises written and released will have an impact on curriculum development.

Good sampling procedures may give insight into the status of knowledge, understandings, skills, and attitudes of students in a particular target population. This can promote curriculum improvement and innovation. Problem areas in the curriculum may be

identified. From the National Assessment, there have been some problem areas identified in the Citizenship results. On an exercise dealing with freedom of speech, a large percentage of 13, 17, and adult age groups indicated that they would not allow sample controversial statements to be made on radio or TV.[14] This showed a lack of understanding or valuing of the Constitutional right of freedom to express controversial or unpopular opinions.

The results on the Citizenship assessment indicated that black, urban students in our large cities compared poorly on knowledge about the structure and function of government to the national average performance on the same exercises.[15]

On the other hand, there are potential difficulties with assessment data which represent national levels of performance. Even though the data were not collected with this intention and were reported in census-like form, the results of National Assessment are being treated like national norms. Several states have conducted their own assessments duplicating the National Assessment model so that they can make direct comparisons between their state results and the various national, regional, and subgroup results. There is the potential of great mischief in this approach,

---

[14] Campbell, Vincent N., et al, Report 2, Citizenship: National Results, Education Commission of the States, National Assessment of Educational Progress, Denver, Colo., November, 1970, p. 34.

[15] Norris, Eleanor L., Vincent N. Campbell, Manford J. Ferris, and Carmen J. Finley, National Assessment Report 9, 1969-1970 Citizenship: Group Results for Parental Education, Color, Size, and Type of Community, Education Commission of the States, National Assessment of Educational Progress, Denver, Colo., May, 1972, pp. 63-65.

for it may lead to unfair comparisons between groups, states, and regions. In the assessment reports of some states, tables of percentages have been presented without any interpretation or explanation. Some school systems have been presented in a very bad way without any reference being made to the kinds of variables involved in the different learning situations. Such variables as per pupil expenditures, educational level of parents, and motivation of pupils do have an impact on the learning situation. These and other variables cannot be ignored in interpreting the results of assessment.

Here, it is not being suggested that assessments should not be conducted because there are potential misuses of the data, but it is being pointed out that misuses of data do occur. This does mean that those engaged in assessment at national, state, and local levels have the responsibility to be constantly engaged in an educational program to aid those using the data to make correct interpretations of it. We need these kinds of information for decision making, but if the data are misused or misinterpreted, then the decisions based on them may not be good ones.

Finally, where accountability is being applied to a total organization such as a school, a district, or a state, the National Assessment model may be used with little or no modification. It was designed to accurately establish what the level of performance on a given set of objectives was in a population, and it can be used to do this for accountability purposes as well as assessment purposes. Likewise, it can assess subgroups of the population and identify specific strengths or weaknesses in the performance of a

given subgroup. The model is an excellent instrument for carrying out accountability in this kind of situation.

SUMMARY

In closing, there are several summary statements which can be made. The National Assessment model is no doubt the best and most comprehensive procedure that has been designed for collecting data for these purposes. The model can and has been successfully adapted for use at the state and local levels, and as states become more active in assessment, there will be many adaptations of it made to fit local needs. Finally, where the concern is with the evaluation of group performance, the model may be used for accountability purposes. With the current rapid development of state assessment and accountability programs, it is expected that the model will be widely used for these purposes in the coming years.